. 484
959

Williams, Lubin, and Goodnow

ol. 73
o. 14

Psychological Monographs

General and Applied

Impaired Performance with Acute Sleep Loss

By

Harold L. Williams, Ardie Lubin, and Jacqueline J. Goodnow

Walter Reed Army Institute of Research

Price $1.00

Edited by Norman L. Munn

Published by the American Psychological Association, Inc.

Psychological Monographs:

General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology* with the *Psychological Monographs*

Manuscripts and correspondence on editorial matters should be sent to the Editor. *Psychological Monographs* publishes comprehensive experimental investigations and other psychological studies which do not lend themselves to adequate presentation as journal articles. Major space is given to the author's original contribution; introductory and bibliographic materials, as well as statistical tables and graphs, must be kept within reasonable bounds. Tables, graphs, and appendix materials which deal with detail not essential to adequate presentation of the findings may be made available through the American Documentation Institute—for details of this procedure, see the APA *Publication Manual.* Preparation of manuscripts for publication as monographs should follow the procedure given in the APA *Publication Manual.* Publication in *Psychological Monographs* is free of cost to the author, except in cases where early publication is requested or author's alterations are made in galley proofs.

Correspondence on business matters should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St., N.W., Washington 6, D.C. Address changes must arrive by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee third-class forwarding postage.

Vol. 73, No. 14 Whole No. 484, 1959

Psychological Monographs: General and Applied

IMPAIRED PERFORMANCE WITH ACUTE SLEEP LOSS[1]

HAROLD L. WILLIAMS, ARDIE LUBIN, AND JACQUELINE J. GOODNOW

Walter Reed Army Institute of Research

THE INITIAL PROBLEM in studying sleep loss is the difficulty of finding any performance, either psychological or physiological, which is consistently impaired. More than 60 years after the first experiment in 1896 (Patrick & Gilbert), a recent review of performance on motor and mental tests concluded that "subjective attitude" (appearance, mood, and behavior) is the only factor seriously affected by sleep loss (Tufts College, Institute for Applied Experimental Psychology, 1949, p. 1). For physiological performance, most of the functions tested show a similar lack of impairment (Kleitman, 1939; Lee & Kleitman, 1923; Tyler, Goodman, & Rothman, 1947) and the one measure which has shown a consistent effect—changes in alpha rhythm (Bjerner, 1949; Malmo, 1958; Tyler et al., 1947)—presents a number of apparent inconsistencies in its relationship to psychological performance.

In response to these difficulties, the most frequent proposal is that the appropriate task or function has not yet been found. For psychological performance, for example, it has been argued that tests should be nonmotor (Weiskotten & Ferguson, 1930) more complex (Laslett, 1924), more demanding of higher mental activity (Kleitman, 1939), or, at the other end of the scale, less challenging (Wilkinson, 1957), less motivating (Ax, Fordyce, Loovas, Meredith, Pirojnikoff, Shmavonion, & Wendahl, 1957), or longer and less amenable to compensatory effort (Edwards, 1941). This search for a sensitive task was not rewarding either in past studies or in our own first efforts. We have come to adopt instead an approach found in studies of fatigue (Bartley, 1957; Bartley & Chute, 1947; Welford, 1953). With this approach, the most general problem is to identify and measure the sensitive aspect of performance on any task. The more specific problems, together with the sources from which we have borrowed some solutions, may be summarized as follows:

1. What is the over-all nature of impairment? The recent emphasis in studies of fatigue has been on the increasing unevenness of performance rather than on any gradual and continuous decline. There are several ways of viewing unevenness. We have followed Bills (1931) in regarding it as the result of blocks or brief periods of no response which increase in frequency and duration with continued mental work, performance between blocks being maintained at the initial level or close to it.

The idea of impairment as brief periods of no response has been extended to sleep loss on two previous occasions, both times with profit. Warren and Clark (1937) used Bills' definition of a block as a pause in responses that was more than twice as long as the average response time. They found that the number of blocks showed a decided increase with sleep loss, while the customary measures of modal or average response gave the usual result of little or no effect. The concept of blocks was taken up again in an excellent but little-known study by Bjerner (1949). Bjerner used the some-

[1] This monograph is part of a two-year study of the effects of sleep loss on performance, personality, and social interaction. Our debts are numerous. To single out only some of them, we are indebted to Charles Gieseking, Irvin Rubinstein, and Betty Shanks for a major share of test administration; to Ethel Eldridge and Ometta Kearney for both test administration and data analysis; to Herman Becker and Pearl W. Mack, who, together with members of the hospital nursing staff, were invaluable in their constant care of *Ss*; and to David McK. Rioch, George Crampton, John Armington, and Edward Murray for constructive comment and practical assistance.

what more general term "lapses" instead of "blocks," and found that lapses in the form of long reaction times were significantly associated with changes in alpha rhythm and pulse rate. Largely on this basis, Bjerner regarded lapses during sleep loss as reflecting brief periods of sleep or a condition like sleep, an approach which we have followed in the present study.

2. In what aspect of performance will impairment appear? In the studies so far cited, lapses were related to changes in speed and only indirectly to the occurrence of errors. In the work of Broadbent (1953), however, both types of change were tied to the way in which lapses affect performance on different types of task. As a rule, tasks are grouped in terms of the common mental process which they are thought to represent—all learning tasks, for example. In contrast, Broadbent's grouping is in terms of whether tasks are "paced" or "unpaced." Unpaced tasks are those like reaction time, where the stimulus remains present until the *S* makes a response, or like most problem-solving tasks where the *S* can determine when he will proceed from one part of the problem to another. In all of these tasks, a response can be deferred until a lapse is over, so that impairment will appear as a change in speed or output but not in accuracy. On paced tasks, however, the stimulus is present for only a short time and the response must be made within a brief and prescribed interval: a vigilance or a radar task, for example. Under these conditions, a lapse coinciding with the appearance of the stimulus can produce an error, particularly an error of omission.

With one modification, we have adopted Broadbent's division of tasks. We have preferred to group tasks in terms of their being *S*-controlled or *E*-controlled, with the pacing of responses as one critical feature—certainly a major one—that may be controlled either by *S* or by *E*.

3. What conditions promote or modify the appearance of impairment? The negative and inconsistent effects of sleep loss have often been attributed to variations in such conditions as task duration, stimulus monotony, knowledge of results, and the challenge of a task. It is important to separate these conditions from the problem of measuring the sensitive aspect of performance because failure to measure the sensitive aspect can, under any condition, result in a spurious picture of no impairment. In the two kinds of task, *S*-controlled and *E*-controlled, we studied the effects of two of these conditions: (*a*) knowledge of results, and (*b*) uninterrupted and unchanging stimulus and response conditions.

The remainder of this monograph is devoted to exploring these general ideas and to determining how well they account for the apparently inconsistent effects of sleep loss. We hope to state them fully enough so that they can be used as a guide in testing for impairment and as a starting point for relating changes in test performance to physiological changes, notably to changes in the EEG alpha rhythm.

Procedure and Design

The results are drawn from two studies of sleep loss, one in 1956 and the other in 1957. There are some differences between these two in the sample of *S*s and in the design. In both studies, however, the experimental setting and general procedure reflect our aim of providing a noncoercive, nonrestrictive atmosphere with a minimum of heavy physical work.

Design. Both studies were designed so that the *S*'s performance could always be evaluated against two standards: (*a*) the *S*'s own performance before sleep loss and after recovery, and (*b*) the performance of control *S*s who were given the same tests but allowed to sleep 8 hr. each day.

The general plan for both studies covered an 11-day cycle and is given in Table 1. The major design differences between 1956 and 1957 were two. In 1956, the sleep loss goal was 72 hr.; in 1957, 98 hr. In 1956, all *S*s were run through two 11-day cycles; the experimental *S*s in the first cycle became controls in the second and vice versa. This was not satisfactory, as experimental *S*s tended to lose interest after what they deemed their real job was completed. Accordingly, in 1957, *S*s went through only one cycle and the next cycle started with a fresh batch of *S*s.

At the beginning of a cycle, a group of 10 to 12 men was divided into experimentals and controls on the basis of one or two days of screening tests. In 1956 the division was based on a sociometric variable. In 1957, two *S*s who met an EEG

TABLE 1

General Design for Sleep Deprivation Experiments

Group	Period		
	Baseline	Sleep Loss	Recovery
	Days 1–4	Days 5–8	Days 9–11
Experimental	Normal wake-sleep pattern	74-98 hr. sleep loss	Normal wake-sleep pattern
Control	Normal wake-sleep pattern	Normal wake-sleep pattern	Normal wake-sleep pattern

criterion were selected as experimentals. The remaining *S*s were matched as far as possible on the basis of scores on the Army Classification Battery (ACB) and then assigned randomly to experimental or control groups. The 1956 *S*s were not told into which group they fell until the end of the baseline period. They could, however, easily infer this from the location of their sleeping quarters and in 1957 we made this information more formal by posting the division of *S*s on the second baseline day.

Procedure. To avoid diurnal variations, any particular test was given at the same time of day for each testing, and the time of testing for experiment and control *S*s was distributed as evenly as possible. The major exception was one occasion in the 1956 study when the experimental *S*s were tested in the early hours of the morning after approximately 70 hr. of sleep loss.

Because of time limitations for both *E* and *S*, not all *S*s were given all tests. With the exception of the selection of *S*s for EEG testing, *S*s were assigned to tests on the basis of availability and, to a lesser extent, ACB scores.

Subjects. In both 1956 and 1957, the *S*s were volunteers from groups of army enlisted men. In 1956, 24 *S*s came from a group of religious conscientious objectors. In 1957, 50 *S*s came from a regular field hospital unit stationed at Fort Belvoir, Virginia. The age range for the 1956 group was 19–22 yr.; for the 1957 group, 18–45 yr. Since one might expect differences between the performance of these two samples, all results presented will be labeled according to the year in which they were gathered.

Experimental setting. The experimental ward was a self-contained unit that consisted of dayrooms, a nurse's office, and sleeping quarters for the experimental *S*s. Control *S*s spent their days on this ward, but slept in a separate ward. All testing was conducted on two wards adjoining the experimental one.

The dayrooms were equipped with television, radio, record players, table tennis, and a variety of craft and hobby kits. Extra food was always available. In 1957 coffee was also available. (Because of their religious views, the 1956 *S*s did not drink coffee.)

A 24-hr. watch was maintained to prevent *S*s from sleeping or napping. Part of this watch was kept by a staff of three nurses and six corpsmen who worked on eight-hour shifts. In addition to these, project administrative officers and control *S*s were often available. The emphasis throughout was on persuasion. If an *S* dozed, he was awakened immediately, and, outside of testing sessions, was taken for a walk, engaged in games, given a shower, etc. As it turned out, much of this work was taken over by the *S*s helping each other and trying to avoid the intervention of "assistants" who were not from their own unit. On the ward, this was made easier by the concentration of *S*s in a limited number of rooms. In the testing rooms, and on the way to and from these, either *E* or an assistant was continually present.

Motivation. The following details are given in view of the emphasis on motivation as a major determinant of whether or not impairment appears (Ax et al., 1957; Wilkinson, 1957).

The *S*s received, in addition to their regular base pay, the relatively small amount of twenty-five dollars. Most of their interest, however, came from competing with one another and from turning the ability to endure sleeplessness into a question of "guts." This picture of a general high level of motivation is supported by three points:

1. All *S*s completed the task; and in 1956, a number of *S*s requested that they be allowed to see how long they could hold out, a request we acceded to by allowing some of them to continue for 89 hr. (the initial goal was 72 hr.). This is in contrast to Tyler's (1946) study where 15% of the *S*s dropped out before the end of 56 hours.

2. The *S*s denied their sleepiness. On scales for both sleepiness and fatigue, the 1957 *S*s rated themselves as more fatigued than sleepy, and gave higher ratings of sleepiness to their fellows than to themselves. The self-ratings of sleepiness were

considerably below the ratings given by observers (Murray, Williams, & Lubin, 1958).

3. Incidents of anger or irritability were, to our surprise, rare. Instead, the *S*s' moods were most often marked by listlessness and a sense of simply putting one foot after the other in a dogged attempt to keep going. These results are similar to those reported by Eagles, Halliday and Redfearn (1953) for a pair of cooperative and gently handled *S*s.

Methods of analysis. Whenever possible, the average Spearman rank-order correlation with hours of sleep loss was computed. The a priori rank order was determined by using the hypothesis that impairment would increase as hours of sleep loss increased. The distribution of this rank-order coefficient, which we have called K, is described by Lyerly (1952). Occasionally, to measure the size of change regardless of trend, the measure K was supplemented by pooling baseline and recovery results and comparing them to the sleep-loss period. An average Kendall tau rank-order coefficient was computed. The distribution of this coefficient, which we call T, is described by Jonckheere (1954). For all results, an asterisk after a value of K or T refers to values at or beyond the .05 level of confidence.

Hours of sleep loss or wakefulness were calculated from 0600 on the day which ended in the *S*'s first sleepless night. For recovery testing, the reference R_1 throughout refers to the day after the *S*'s first sleep. As is typical in experiments on sleep loss, this sleep was little different in length from the *S*'s usual number of hours of sleep.

Subject-Paced Tasks[2]

In a task which is completely *S*-controlled, the *S* determines the time when the stimulus appears, the duration of the stimulus, the time within which the response must be made, the distribution of work and the acceptable level of performance.

In the present section we shall be mostly concerned with tasks in which the effective response time and, indirectly, the duration of the stimulus are *S*-controlled. As a general term, we shall refer to these tasks as *S*-paced. On all of them, we expect to find changes in speed.

One of these *S*-paced tasks is simple reaction time where the lapse hypothesis allows us to predict a particular form for the changes in reaction time; namely changes in the range and distribution of response times. Three of the remaining tasks, like addition or concept attainment, illustrate an impairment in speed but not in accuracy. And the final task points to changes in the quality of performance that occur when the *S* can control the standards for performance.

Reaction-Time Tasks

Typically, a reaction-time test consists of a warning signal and a foreperiod followed by a signal to which the *S* responds as rapidly as possible. The *E* controls the time when the warning appears, the length of the foreperiod, and the time when the signal appears. The duration of the signal, however, and the effective response time are *S*-controlled. The signal generally remains present until the *S* responds. The time allowed for response is either infinite or, if controlled, so generous that it extends well beyond the time *S* takes for any response. On a task like this, the usefulness of the lapse hypothesis lies in predicting the form taken by changes in reaction time.

The Nature of Decrement

Suppose that lapses with sleep loss show the same increase in frequency and duration that Bills (1931) reports for prolonged, uninterrupted work. When a lapse coincides with the appearance of the signal, the response will be delayed. When the signal occurs between lapses, the speed of response will be unaffected. There will be then some trials on which reactions will be as fast as any during the baseline period, and some trials on which reactions will be considerably slowed. The difference between the slowest and fastest responses will become progressively greater as lapses increase in duration. As lapses increase in frequency, the distribution of reaction times should become positively skewed. Gradual impairment (the obvious alternative to the lapse hypothesis) implies a constant increase in all reaction times, with no change in the range and distribution of responses. Both

[2] For the administration and analysis of these tests, we are indebted to Ometta Kearney and Ethel Eldridge (reaction-time tasks), Betty Shanks (concept attainment), and Irvin Rubinstein (predicting probable events).

hypotheses imply an increase in mean reaction time.

In all four of the reaction-time tasks we used, the increase in range and in positive skew is the most striking result. There is as well a slight over-all increase in reaction time so that the fastest responses during sleep loss are slightly slower than the fastest in the baseline period.

To illustrate these effects, the results of a reaction-time task, given in 1956 to eight experimental *S*s, are presented below.

Apparatus and procedure. The *S* sat at the center of a table. Before him was a center light (blacked out), two response keys (left and right), and a small loudspeaker. To the extreme left and right of *S* were two lights (each 23 in. from the center of the table). *S* fixated the center spot. A click from the loudspeaker warned *S* that in 2 sec. one of the two sidelights would be turned on. With the appearance of a light, *S*'s task was to press the corresponding key (left or right) as rapidly as possible. Pressing the appropriate key turned off the light. The length of the warning period (2 sec.) was controlled by a Microflex Timer. Reaction time was measured by a Standard Precision Electric Timer (SPET).

Both the Timer and the SPET were operated by *E* from an adjoining room with a thick wall between *E* and *S*. The occasions of testing are shown on Fig. 1. Each testing session covered 72 trials.

Results. Figure 1 shows a marked increase in reaction time during sleep loss, whether we consider the mean or the median. After 78 hr. of sleep loss, mean reaction time is twice as long as on the last baseline day. Statistically, reaction times during the sleep-loss period are significantly longer than during the baseline and recovery period (tau [T] for mean reaction time is .79 for the mean and .60 for the median). Also, increases in reaction time are strongly related to increasing hours of sleep loss (K for the mean was .82, for the median, .85).

Of most interest to us at the present moment is the way in which various trials contribute to these general measures (mean or median). In Fig. 2 we present the average of the 10 shortest and 10 longest reaction times through the baseline, deprivation, and recovery periods. Figure 2 leaves us in no doubt that *S*s during sleep loss are capable on some trials of coming very close to their best performance during the baseline period. The 10 best trials during sleep loss show only a slight increase over those during the baseline period. The increase, however, is consistent (T = .65, K = .85). Far more striking is the enormous increase in the duration of the long reaction times. The *S*'s poorest performances become progressively worse, even though his best performances remain close to his original level.

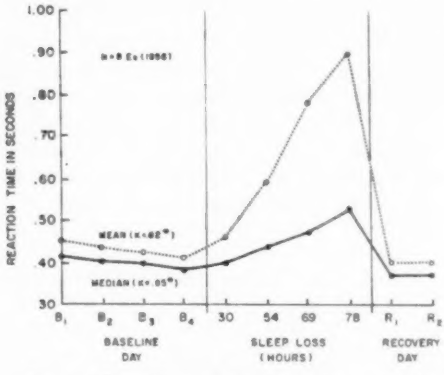

FIG. 1. The effect of sleep loss on a reaction-time task. [United States Army photograph]

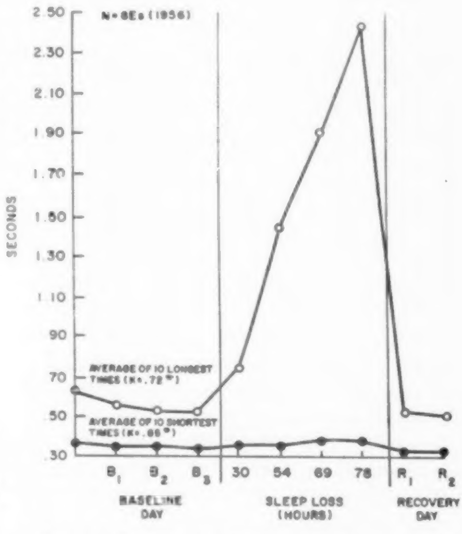

FIG. 2. Changes in extreme reaction-time with sleep loss. [United States Army photograph]

This increase in range of reaction times can be regarded as support for the argument that with sleep loss lapses increase in duration. Do they also increase in frequency, as Bills (1931) found for fatigue? If we apply Bills' definition, a "lapse" or "block" would be any reaction time that is twice as long as *S*'s average reaction time on the last baseline day. Figure 3 shows that these lapses occur more and more often as sleep loss increases.

These increases in the duration and the frequency of lapses are found consistently on all four of the reaction-time tasks that we used.

Task Conditions Affecting Decrement

Task duration. The four reaction-time tasks we used varied in several ways: the number of trials given, the regularity of the foreperiod, and the duration of the foreperiod. When the experiment was designed, our concern was predominantly with the latter two conditions.[3] As it turned out, the reaction-time data, when analyzed with respect to the regularity and duration of the foreperiod, gave inconsistent results. In contrast, we consistently found that reaction time was an increasing monotonic function of task duration (a variable which was not directly considered in the experimental design).

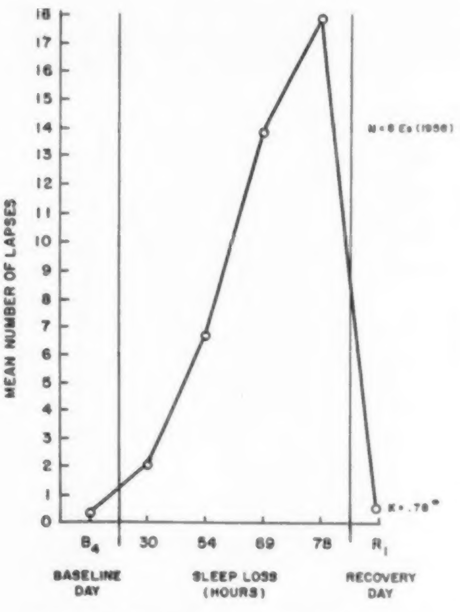

FIG. 3. Incidence of long reaction-times (twice the baseline average). [United States Army photograph]

Table 2 presents a summary of results and conditions in the four tasks. Task A, which ran for 72 trials without interruption, was highly sensitive as a whole to sleep loss. In the other four tasks (B, C, D, and E), there was a 1-min. break between conditions and some changes in conditions after the break. Here there was a tendency for only the last part of the task to be sensitive, regardless of the variations in other conditions.

The critical factor appears to be the length of time, without interruption, that the same stimuli occur and the same response is required. Breaks in a task appear to delay the occurrence of lapses until the later stages of the task. From our data it is not clear, however, whether the time break derives its effectiveness from being simply a rest period or from introducing a welcome change into the stimulus situation. The greater importance of the latter is certainly implied by the emphasis on stimulus and task monotony in a number of studies of fatigue (Broadbent, 1953a; Mackworth, 1950; Solomon, 1948), and it fits with Wilkinson's (1958) emphasis on the sensitivity, to sleep loss, of "predictable" tasks.[4]

Knowledge of results. In Task C, *S* was given knowledge of results. The *S* faced SPET. For half the trials, this timer was in circuit so that it registered, within *S*'s view, the reaction time for each trial. The *S* reset the timer after each response. For

[3] At the time, we were following previous work (Rodnick & Shakow, 1940) that had emphasized the relevance of these variables to the *S*'s ability to maintain a set.

[4] More detailed material on the effects of task interruptions is presented in the section on *E*-paced tasks.

TABLE 2

DESCRIPTION OF REACTION-TIME TASKS

Task	Ss	Stimulus (St) and Response (R)	Order and Nature of Conditions				Results	
			Foreperiod	Knowledge of Results	Trials for Each Condition	Approx. Task Time	Tau	K
A	1956 Es–8	St—1 of 2 lights R—Press 1 of 2 keys	Constant (2 sec.)	No	72	10′	.79[e]	.82[e]
B	1956 Es–10	St—1 light R—Press key	1. Irregular[a] 2. Regular[b]	No No	50 50	30′ 30′	.52[e] .63[e]	.40[d] .95[e]
C	1957 Es–8 Cs–8	St—1 light R—Release key	1. Regular 2. Irregular 3. Regular 4. Irregular	No No Yes[c] Yes	25 25 25 25	30′ 30′ 30′ 30′	.23 .20 .12 .62[e]	.25 .12 .62[e] .81[e]
D	1957 Es–4 Cs–4	Same as C	1. Regular 2. Irregular	No No	50 50	30′ 30′	.50[d] .80[e]	.50 .50
E	1957 Es–6 Cs–6	Same as C	1. Irregular 2. Regular	No No	50 50	30′ 30′	.53[e] .84[e]	.75[e] .83[e]

[a] Intervals were either 5″, 1″, 15″, 2″, or 25″, in random order.
[b] Interval constant within each block of 10 trials; for successive blocks, 5″, 1″, 15″, 2″, 25″. *S* is told that after each block the interval will change.
[c] Timer registers *S*'s reaction time within *S*'s view. *S* resets timer after each response.
[d] Significant at .05 level.
[e] Significant at .01 level.

the other half of the trials, the clock was not in circuit. The instructions about the use of the timer were simple: "Whenever you finish lifting your hand off the key, look at this clock on the left. That tells you your reaction time. Fifty means you took one-half second; 30 means you took three-tenths of a second, and so on. Before putting your hand back on the key, reset the clock." These instructions leave a great deal up to *S* and some of the *S*s regarded resetting the clock as a purely motor part of the task and a distraction to boot. In this particular form, knowledge of results did not help the *S* to resist the effects of sleepiness (see Table 1).

This result was surprising, in view of the emphasis placed on knowledge of results by, for example, Mackworth (1950) in his fatigue studies and by Wilkinson (1957) in his studies of sleep loss. The source of this discrepancy may lie in the way in which "knowledge of results" was given. In Mackworth's task, for example, the procedure was to say to *S* after each trial either "Yes, that's right," or "You missed that one." In our Task C, results were available but were not labeled by *E* or by any other standard as success or failure. If the *S* did attend to the timer, he had to define for himself what constituted failure and success; both his standards and his perception of the situation could readily change.

Summary of reaction-time tasks. Sleep loss persistently produced an increase in reaction time. On all four tasks, this increase took the form of a progressively greater unevenness of performance, with an increase in both the range and the positive skew in the distribution of reaction times. These results follow from and agree with the hypothesis that sleep loss produces an

increase in the frequency and duration of lapses. In contrast, there may be a gradual deterioration affecting every reaction time, but it is very slight.

The major condition promoting impairment is task duration, particularly the length of time that a task continues without interruption, and without a change in either the stimulus conditions or in the response required.

Tasks in Which Speed and Accuracy Can Be Measured Simultaneously

In a simple reaction-time task, speed is usually the only appropriate measure of performance. In more complex tasks where *S*, for example, has to discriminate between different stimuli, has to combine information, or make inferences, there are usually several ways of measuring efficiency. Since these measures may be relatively independent of one another, it is possible for one aspect of performance to decline while another remains stable, or even improves. The two aspects usually measured are speed and accuracy.

How does it follow from the lapse hypothesis that, in *S*-paced tasks, accuracy will be less sensitive to sleep loss than speed? Three assumptions are involved: (*a*) Between lapses, performance can be maintained at close to the normal level. At this level, lapses simply represent "time-out" between responses. (*b*) If a lapse interrupts a chain of operations so that the *S* loses his place or makes an error, the *S* can use his unlimited response time to start again or to correct the error. (*c*) In most tasks where both speed and accuracy are measured, the task-set is such that the *S* will sacrifice speed for accuracy.

To determine the usefulness of these assumptions and to check the general expectation of an impairment in speed rather than accuracy, we present below the results of three *S*-paced tasks: adding, concept attainment, and sending instructions.

An adding task. The main instructions were: "Here is a list of numbers (one-digit). I want you to add each two numbers and write the answer here. . . . You will be doing this for three minutes. . . . Just keep going, as fast, but as accurately as you can." The test was given both in the morning and evening, and Fig. 4 shows the results for evening testing. Even after 86 hr.

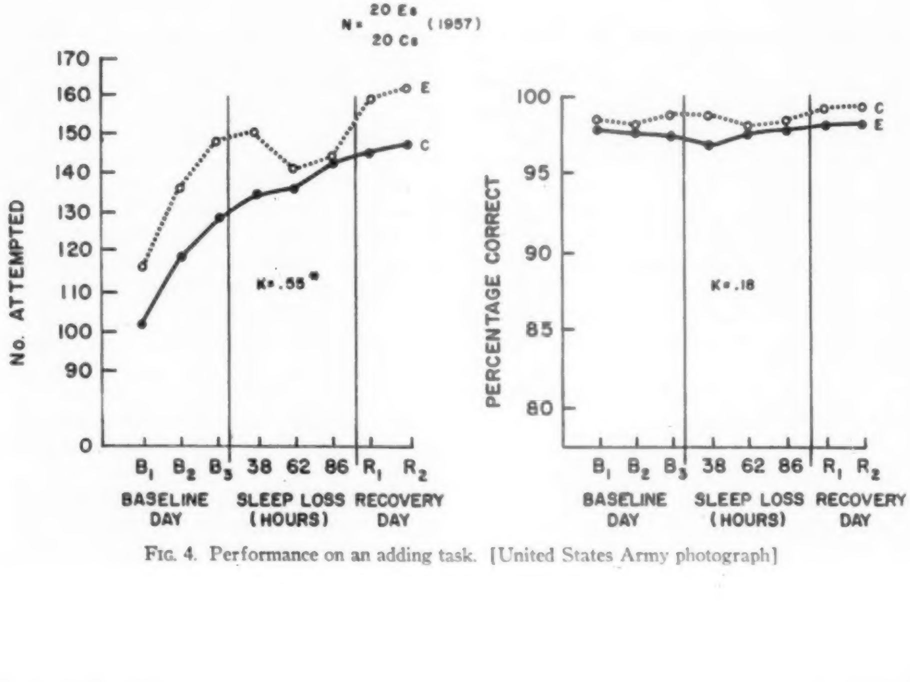

FIG. 4. Performance on an adding task. [United States Army photograph]

of sleep loss, the percentage correct shows minimal change. The number of pairs attempted, however, decreases significantly after 48 hr. of sleep loss.[5]

A communication task. In this task, designed and already reported by Schein (1957), the *S* sent and received instructions about the placement of domino-like pieces of cardboard into a pattern (10 or 25 pieces to a pattern). The sending part of the task was essentially *S*-paced. The *S* sat before a pattern of pieces. His task was to tape-record instructions that would allow someone else to build the pattern. The *S*s had already become practiced in a "code" for giving these instructions, describing each piece as placed either horizontally or vertically, and specifying its point of contact with some previously placed piece. The *S*s were told that they could work as fast as they wanted to but that they would be scored on accuracy; if they made an error or omitted something, they should correct it or fill in the necessary information.

In this task, the number of pieces described incorrectly shows an insignificant increase with sleep loss. In contrast, the time taken for sending shows a progressive, significant increase with sleep loss. Also the number of errors made, but corrected, increases significantly with sleep loss: At 70 hr. of sleep loss, it is more than double the baseline period.

A concept-attainment task. This task, which is described fully in Bruner, Goodnow, and Austin (1956), requires the *S* to determine the properties that define class membership. The *S* faces an array of 81 cards on a board and these cards differ from one another in four respects: the figures on them vary in their shape (square, circle, or cross); their color (red, green, or black), their number (1, 2, or 3), and the lines bordering the card can be 1, 2, or 3 in number. The *S* is told that these cards can be grouped in different ways; e.g., all cards with two green crosses belong in a group, all other cards do not belong. The *E* then explains that he has a grouping in mind, and points out one card as belonging to this group. The *S* proceeds by choosing other cards one at a time. After each card, *S* is told whether or not the card belongs in the group and he continues until he can state what properties a card has to have in order to be in the group.

The *S*s were given six problems on each day of testing, with 5 min. per problem (a generous time limit). If the problem was not solved at the end of this time, *S* was told the answer, and then went on to the next problem.

Under these conditions, we find again no significant change in the general quality of performance. There is a slight decrease in the number of problems solved correctly (K is .12) and a small increase in the number of choices per solved problem (K is .22). In contrast, the time per solved problem shows a significant increase (K is .42). A general picture of results is presented in Fig. 5. All but two *S*s (out of twelve)[6] showed an increase in time, and after 77–81 hr. of sleep loss, these increases range from 16% to 71%. In contrast, all but one of the control *S*s required, on the same testing day, less time than they required at the end of the baseline session.

In short, on this task *S*s can again maintain their previous level of correctness, or a level very close to it, but they need more time.

Summary of complex S-paced tasks. On three tasks where it was possible to measure both speed and accuracy, changes in time but not in accuracy were consistently found. These tasks—addition, sending instructions, and concept attainment—differ in their

[5] Morning testing shows a similar slowing of work during sleep loss; K is .42 (.01 level), and upon the first recovery testing, the mean number attempted rises from 130 to 150. Percentage correct also shows some decline with sleep loss; K is .32 (.05 level), but the size of the change is small (from a baseline of 98.9% to 97.9% after 72 hr.). All results are drawn from the work of Nathene Loveland.

[6] In all, there were 16 experimental and 16 control *S*s, but time scores were not recorded for the first 8 *S*s.

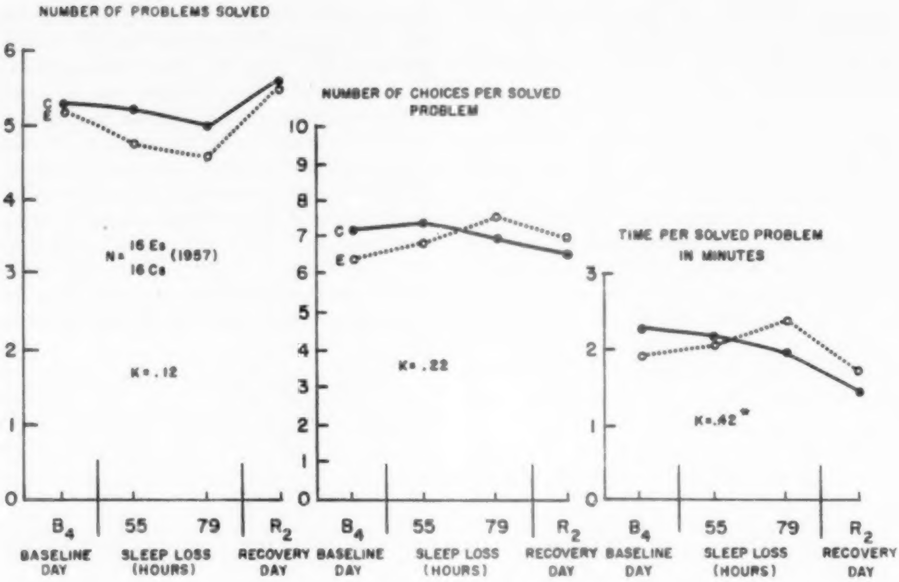

FIG. 5. Performance on a concept-attainment task. [United States Army photograph]

length and complexity but share two common properties. They are *S*-paced and the stress is on accuracy.

Changes in the Standard of Performance

In the previous three tasks, changes in speed can be expected simply on the grounds that lapses or brief periods of no response will occur. The lack of change in the quality of performance depends upon *S* having both the opportunity and the incentive to maintain accuracy by correcting errors or by keeping past events firmly in mind.

Suppose that this incentive was reduced or that the goal of the task could be changed by the *S*. In effect, the task would provide more than one level of success, and the *S* would control the level regarded as acceptable. A task in which the *S* predicts probable events provides for such a departure from the usual arrangement of either complete success or complete failure in solving a problem, and performance on it shows a decline in quality.

The *S* attempts to predict as accurately as possible whether a left- or a right-hand key will deliver a chip from a modified "slot-machine" (Goodnow, 1955). Success is set in terms of winning "around 50% or better" of the time. The *S* is playing against a random sequence of 144 events with the machine set to pay off 50% of the time on left and 50% of the time on the right. The *S* perceives the series as forming a pattern that is either very complex or subject to frequent change.[7] The variety of hypotheses about the pattern and the variety of attempts to "beat it" are limited only by the goal the *S* sets himself and his readiness or ability to keep in mind the past events that will reveal the pattern.

The major performance measure is the predictability of a choice from the outcome of the previous trial. This predictability is a measure of the variety of run-hypotheses that an *S* attempts (Hyman & Jenkin,

[7] This perception of a pattern is almost universal on a two-choice task, even when not fostered by *E* (Goodnow, 1955).

1956). Its range is from 0 to 1.[8] With a low predictability score like .04, for example, we find a statement like this:

"I'd usually stay with it if it had won and go over to the other side if it had lost. But it would all depend on how things had been going. Say there's been a lot of long runs and I figured it was time for some little ones. Then, I'd drop my usual system and switch around a bit more. If I figured it was going to go back and forth, I wouldn't go over to the other side if I lost. I'd wait for it to come up."

In contrast is the same *S*'s later record and description of his procedure after 28 hr. of sleep deprivation:

"If it was winning, I'd try it for about four times. If it didn't win, I'd go over to the other side most of the time. It's not a very good system. If I was feeling good, I'd find a way to outsmart it. I know it's not as simple as that." Here, the *S* adopted a simple rule of thumb without the "if's" and "buts" about a variety of past events. His predictability score was .56.

Figure 6 shows how this predictability measure increases during the sleep-loss period and drops again on recovery. The increase appears early in the sleep-loss period, and it is shown by nine of the ten *S*s.

Summary and Comment on S-Paced Tasks

On a variety of tasks in which the *S* controls the time when a response is made or the interval between responses, the consistent result is a change in the speed or rate of performance, but not in its accuracy. These changes in speed stem primarily from an increase in the frequency and duration of lapses. They are also affected by the same task properties that give rise to little or no change in accuracy: the unlimited response time which allows for the delay of a response or for the correction of errors, and the usual orientation toward correctness which prompts the *S* to sacrifice speed for accuracy.

The selection of any particular measure of time depends upon the task structure. In the simple reaction-time tasks that we have

[8] The score is the deviation from .50 of the proportion of times a correct choice is repeated, plus the deviation from .50 of the proportion of times that an incorrect choice is repeated. Additional details on results, apparatus, and procedure are given in Goodnow, Rubenstein, and Shanks (in press).

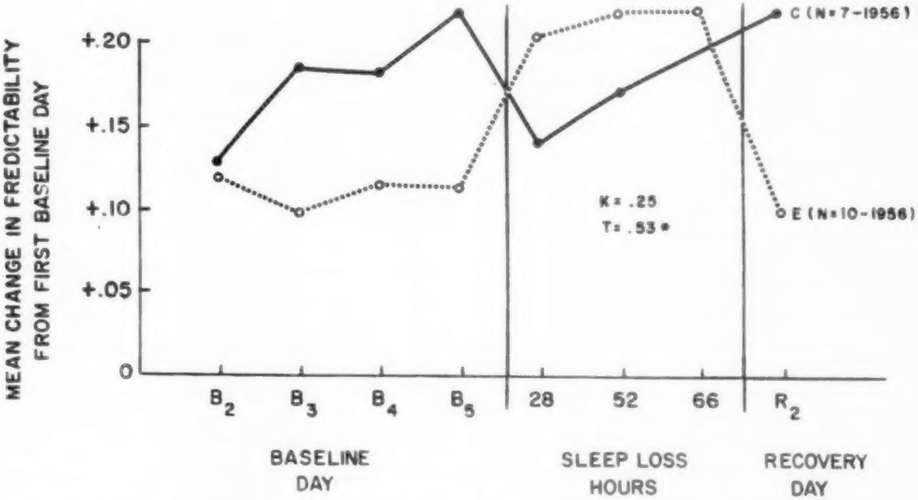

FIG. 6. Predictability of response on a two-choice task. [United States Army photograph]

used, where *E* controls the time when the stimulus appears, a regular reaction-time measure is appropriate, with the emphasis on changes in the range and distribution of these. In a serial reaction-time task, where the *S* determines the time when the next stimulus appears, delays between responses appear to be the most appropriate (Bjerner, 1949; Wilkinson, 1958). On more complex tasks like problem solving, the measures may be over-all time, output, corrected errors, bursts of work, or distribution of work over time, depending upon the way in which performance on the task is affected by lapses and the opportunities the *S* has to compensate for lapses.

All of these suggested measures relate to time, and all of them are proposed for tasks where the *S* controls the pacing of responses. Changes on *S*-controlled tasks are not, however, necessarily limited to time. As Welford's (1953) analysis of fatigue indicates, errors may appear on *S*-paced tasks as a secondary form of deficit, reflecting perhaps the *S*'s anxiety about performance or his growing reluctance to make the effort demanded. Furthermore, changes in the quality of performance appear when the *S* controls the standards of performance. We observed, for example, a decline in the quality of performance on a task which provided an opportunity for the *S* to change his standards of performance where such a shift carried no obvious reminder of failure.

In general, the results from these *S*-controlled tasks strongly suggest that performance will be impaired or changed in one way or another during sleep loss. The aspect of performance affected will depend upon the way in which the task is structured, upon the places where the task allows for a relaxation of performance, and upon the opportunities and incentives to recover from or compensate for the occurrence of lapses. To complete this argument, we turn now to tasks with a completely different structure: tasks where such properties as the duration of the stimulus and the effective response time are *E*-controlled and where the lapse hypothesis would lead us to expect impairment in the form of errors.

Experimenter-Paced Tasks

In most *E*-paced tasks, the time available for an effective response is determined by *E* and is severely limited. The *E* may also control when the stimulus appears, the duration of the stimulus, or the required rate of performance. As in *S*-paced tasks, the response may be simple: in "vigilance" or "watch-keeping" tasks, for example, the *S* is required only to detect the occurrence of a critical signal. In other tasks, the response may be more complex: the *S* may be required to choose among several responses on the basis of information received or to commit to memory information available only briefly.

In *E*-paced tasks, either simple or complex, the lapse hypothesis leads us to expect an impairment in accuracy with sleep loss. The reasoning is as follows. Sleep loss will produce brief periods of no response in the face of stimuli which normally elicit a response, and these lapses will increase in frequency and duration as sleep loss progresses. In an *E*-paced task, the effective response time is limited. The signal in a vigilance task, for example, occurs briefly and without warning. If a lapse occurs at the time when a response should be made, the result will be an error.

Most often we would expect the errors to be errors of omission; i.e., the *S* simply does not respond to the critical signal. There are occasions, however, when we would expect errors of commission, a faulty response rather than no response. In a vigilance task, for example, the *S* will sometimes realize that he has missed a signal, will try to "catch up" and so give a faulty response on the next stimulus. Errors of omission should still be more frequent on any vigilance task where there is an extremely brief interval between stimuli, little time to develop a feeling of error, and no overt sign to the *S* of a gap in his performance. Errors of commission can be expected to be more important when the stimulus is a long and complex one (instructions, for example). On these tasks a lapse may coincide with only part of the stimulus, and, especially if some response is demanded, the

effect will be a faulty rather than an omitted response.

Vigilance Tasks

In vigilance tasks, the *S* is required to make a simple response (usually pressing or releasing a lever) when and only when a critical signal occurs. In the present study, the signals varied in modality from task to task, and were either visual, auditory, or tactual. The aim of this was to ensure that errors were not simply the result of peripheral sensory difficulty (visual stimuli are especially suspect in this regard), and to check on the possibility that lapses were central in origin.

The visual task used the Continuous Performance Test (CPT) developed by Rosvold, Mirsky, Sarason, Bransome, & Beck (1956.)[9] For the first 10 min. the *S* responds by pressing a lever whenever the letter "X" appears in the viewer. (During the second 10 min. *S* responds to "X" only when it is preceded by "A." The results of the "AX" task will be analyzed in a later section.) In all, the X test covers 620 trials with the letter X occurring on 160 trials. The letters occur in a fixed sequence of 31 letters, in which X appears eight times. This sequence is repeated 20 times at the rate of approximately one letter per second; i.e., a total of approximately 10 min.

The auditory task was designed to duplicate the essential features of the visual task. A taped sequence of 30 letters was played to the S, who responded by pressing the lever when "X" was heard. As in the visual task, the rate was one letter per second and the total time was 10 min. The sequence was repeated 20 times, and the letter X occurred eight times in each sequence.

In the vibratory task,[10] the *S* released a signal key whenever he felt a vibration at either his left wrist or his right temple. All told, stimuli could come from any of six body areas, with vibrators (magnetic earphones) at each temple, wrist or knee. In one form of the task, only one area was stimulated at a time; in another, two areas were stimulated simultaneously. The interval between stimuli was again one second. Each sequence of 52 trials contained 18 critical stimuli and the sequence was repeated 12 times. In all tasks, responses were automatically recorded.

[9] We are grateful to Allen Mirsky for the loan of the apparatus.

[10] The vibratory task was designed and conducted by George H. Crampton. All other tasks in this section, with the exception of Schein's (1957) communication task, were administered by Charles F. Gieseking.

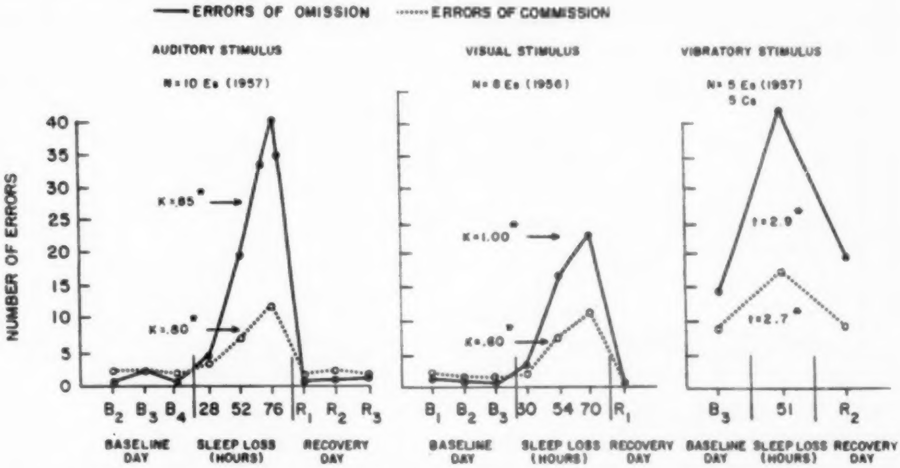

FIG. 7. The incidence of errors on three vigilance tasks. [United States Army photograph]

For illustrative purposes, the detailed results will be given for only one task in each modality. In all three of these tasks the usual condition of no knowledge of results prevailed.

Results. A picture of the size and nature of the decrement is given in Fig. 7 which shows errors of omission and commission for each sense modality. It is of interest that there are consistently more errors of omission than of commission and that omissions are more sensitive to sleep loss. The evidence that lapses occur in all three sense modalities strongly supports the notion that they are central in origin.

Conditions Affecting Decrement

The conditions with which we are concerned are much the same as those discussed in the section on *S*-paced tasks. Task duration is again important and the effect of task interruptions and changes in the stimulus conditions can be shown more clearly. Similarly, the conditions that are roughly called "motivation"—information about results, exhortations to do better, the challenge of a task—appear again to be less effective than task duration.

Task Duration. Figure 8 shows the incidence of errors in 2-min. blocks of the X and AX forms of the Rosvold-Mirsky CPT. On the first day of sleep loss there is little effect, but by about 54 hr. of sleep loss, the second block of 2 min. is producing more errors of omission than the first 2 min. At 70 hr. of sleep loss there is an even steeper rise in errors during the first 6 min., which then levels off.

Figure 8 also shows that the same effect occurs on Form AX, which followed Form X. It is interesting to note that during the first 2 min. of the AX task errors are much less frequent than during the last 2 min. of the X task. This brief recuperative effect may be due to either the short rest pause (1–2 min.) or to a slight change in task (from an "X" to an "AX" task), and the differential effects of these we cannot disentangle with the data we have. Combined, however, they come to play a more and more important part as sleep loss progresses.

As a check on these results with the Visual Vigilance Test, Fig. 9 shows that task duration has roughly the same effect on Forms X and AX of the Auditory Vigilance Test. In both tasks there is the strong implication that the important factor is not simply total time but the duration of an uninterrupted set of stimulus-response conditions. In both tasks there is also a definite interaction of task duration with sleep loss; as hours of sleep loss increase, the detrimental effect of task duration becomes stronger.

Effect of motivation. For both the vibratory and visual vigilance situations, we attempted to vary motivation by varying

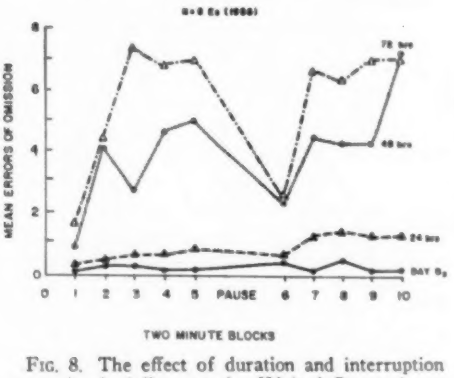

FIG. 8. The effect of duration and interruption on a visual vigilance task. [United States Army photograph]

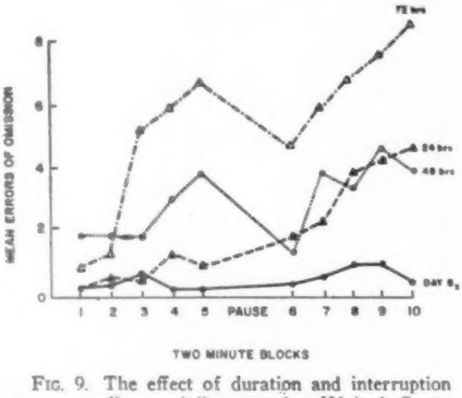

FIG. 9. The effect of duration and interruption on an auditory vigilance task. [United States Army photograph]

knowledge of results. In both tasks it was found that sleep-deprived *S*s with knowledge of results tended to show less impairment than sleep-deprived *S*s without knowledge of results, but the differences were small and somewhat inconsistent.

Effect of motivation on the Vibratory Vigilance Task. For one group of 12 *S*s, only one body area was stimulated at a time. In a second group of 12 *S*s, two body areas were stimulated simultaneously, a condition which makes discrimination of the critical stimuli more difficult. In each group, the six experimental and the six control *S*s were subdivided into high- and low-motivation conditions. Control and experimental *S*s were matched on baseline scores.

The high-motivation groups were exhorted to do better every day. A general knowledge of results was given; i.e., *S*s were not given absolute scores but were told whether they had performed better or worse on the previous test session and the percentage change in terms of their previous score. In contrast, the low-motivation groups were given no knowledge of results and no exhortation.

Difference scores (experimental minus control) were computed for each pair of *S*s on each test session, and the daily averages for the three pairs in each task are shown in Fig. 10. Although the high-motivation group shows less decrement during sleep loss, the differences between the high- and low-motivation groups are significant only for the single stimulus task. The average rank-order correlation, K, is .87 ($p < .01$) for the single stimulus experiment and .27 (n. s.) for the double stimulus experiment.

Effect of motivation on the Visual Vigilance Task, Form 2.[11] On a visual vigilance task, 20 *S*s were again divided into high-

[11] The apparatus for this task was modeled after the Rosvold-Mirsky CPT apparatus. The only differences are: (*a*) *S* released rather than pressed a lever when an X appeared, and (*b*) each letter was illuminated for only the half second during which it passed the aperture.

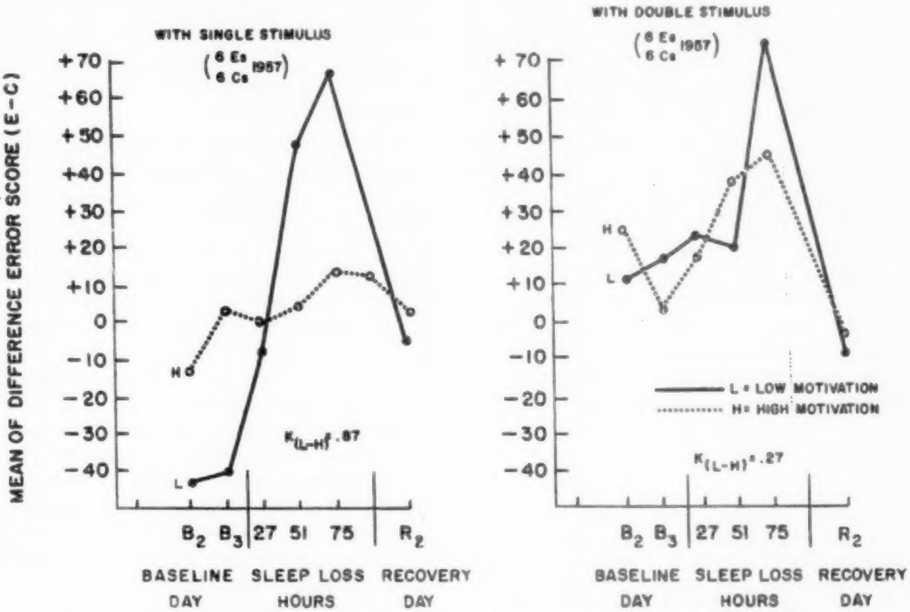

FIG. 10. Errors under high and low motivation on a vibratory vigilance task. [United States Army photograph]

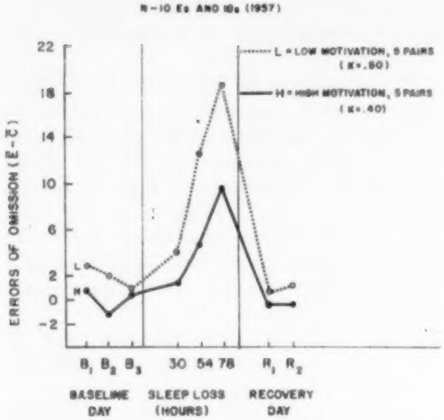

FIG. 11. Errors under high and low motivation on a visual vigilance task. [United States Army photograph]

and low-motivation groups. In the high-motivation condition, *S*s were given their scores at the end of each session, and were encouraged to compete with the other *S*s as well as with their own previous scores. Some information about performance on each trial was also available. An electric counter in the room made a loud click whenever a stimulus was presented, and the click accompanying the X's could be reliably distinguished from the other clicks. This counter was removed from the room for the low-motivation *S*s, who were in addition given no feedback on over-all performance, and were not encouraged to compete.

The results were inconclusive. Figure 11 shows that the low-motivation *S*s appear to be more affected at about 76 hr. of sleep loss, but this is not a significant difference.[12] This result is typical for the procedures we used to vary motivation: namely, the results are in the expected direction, but are not consistently significant.

Summary of vigilance tasks. Sleep loss consistently produced impaired performance on three vigilance tasks. On all three tasks, the major effect was an increase in errors of omission: the *S* missed increasingly often the occurrence of a critical signal. This is the result expected from lapses on a task where the critical stimuli occur briefly and without warning.

Over the three tasks, the modality of the stimulus was varied (stimuli were visual, auditory, or vibratory). The similarity of results on all three tasks makes it very unlikely that lapses are the result of any peripheral difficulty, and suggests instead that they are central in origin.

Of the two task conditions that were considered—task duration and knowledge of results—the former is again the more effective. Errors increase as a task continues without interruption and without a change in the stimulus-response conditions. As sleep loss progresses, errors begin to appear earlier and earlier in the task, and the benefit derived from an initial effort or a break between parts of the task becomes increasingly short-lived.

E-Paced Tasks with a More Complex Response

In all the vigilance tasks we have described, the response required of the *S* is simple—he is to detect the presence of a signal and to indicate this by pressing or releasing a key. In two additional tasks, which are also *E*-paced, the stimulus and response conditions are more complex. In both of these tasks, sleep loss again resulted in an increase of errors.

A communication task. In the previous section on *S*-paced tasks, we discussed the "sending portion of a communication task devised and conducted by Schein (1957). The receiving part of the task is *E*-paced. The *S* sat at a table with 10 domino-shaped pieces. He listened to a tape-recorded, standardized set of instructions which told him how to place the pieces, one at a time, to construct a pattern. The instructions were paced so that about 12 sec. were required to describe the location of a

[12] Neither group, strangely enough, shows a significant correlation of omission errors with hours of sleep loss, perhaps because of the small number of cases. In a third test of the visual X task, however, with eight *E*s and eight *C*s under normal conditions, errors of omission were again significantly related to hours of sleep loss (K = .95*).

single piece and the instructions for each piece were prefaced by the number of the piece. During a given session *S*s completed five problems requiring 10 pieces each.

In this task the number of errors (incorrect placement of pieces) increased significantly with sleep loss. This is in contrast to the results for the sending (*S*-paced) part of the task where the significant decrement occurred in speed but not in accuracy. The errors are errors of commission, which makes sense in view of the possibility of a lapse coinciding with only part of the 12-sec. set of instructions for each piece, and in view of the fact that instructions give the number of each piece so that any gap in response can be detected by the *S* who must fill in the gap in some way before the next piece can be used.

An information-learning task. The stimuli were short items of information, and the response required of the *S* was to commit these to memory. The task was *E*-paced in that the information was available only for a brief period of time. We would expect, with sleep loss, an increasing number of items to be "missed," i.e., not learned to the point of being available for recall. This is, in fact, the major result obtained.

A list of some 428 information items was culled from various sources such as *Information Please Almanac.* From this pool, 275 items were selected. We tried to pick out items that would be of general interest, but not generally known. For example: "Who was the first President to die in office?" By a random procedure, 25 items at a time were selected from the 275 to make up 11 lists, one for each day of the experiment.

On the screening day, each *S* was pretested on the 11 lists so that a baseline score could be obtained. Answers were not given during this pretest. The control *S*s were matched individually with the experimental *S*s on this baseline score.

During each session, while the *S* stood with his eyes closed having his EEG recorded, *E* read aloud the 25 items and their answers. He read at a prescribed rate, pausing 5 sec. between question and answer. The answer was read only once and after a 5-sec. pause, *E* went on to the next question. Ten minutes after the first presentation, the list was given again to test for immediate recall. The immediate recall score with baseline answers subtracted is called R_1. If, on any item, the *S* could not respond or gave an incorrect reply, he was given the correct answer. If he answered correctly, *E* said "Correct." A few minutes later the *S* was again tested for recall, but this time his answers were not corrected. The number of correct answers, after subtracting the baseline correct answers, is designated R_2.

On the following day, some 23 hr. later, the same list was presented a third time to test for delayed recall. (This delayed recall score, with baseline score subtracted, is called R_3.) The *S*'s answers were corrected. Later in that session, a new list was presented, and so on, with a different list each day. On any one day all *S*s received the same list.

Results. Figure 12 shows the results with 10 experimentals and 10 matched controls, for the initial recall score, R_1. This score, R_1, is essentially the equivalent of correct responses on a vigilance task, the score declining as more and more questions are missed (responded to mostly with the answer "Don't know" or "Don't remember"). The sleep-deprived *S*s show impairment at about 28 hr. of sleep loss, and the decline continues throughout the 76 hr. of sleep loss (K = .67*). There is a signifi-

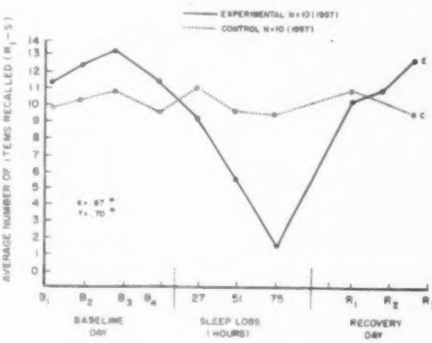

FIG. 12. The effect of sleep loss on immediate recall. [United States Army photograph]

cant recovery immediately after the first sleep. The improvement continues and the scores return to about the control level after 25 hr.[13]

A failure in immediate acquisition or immediate recall is easily accounted for by a lapse hypothesis. The *S* simply fails to take in information that is only briefly available. This failure, however, is not the only important facet of a learning task. For these other facets, the scores R_2 and R_3 are relevant. R_2 (recall a few minutes after R_1) declines but is more resistant to the effects of sleep loss than R_1, which is not surprising since R_2 follows a few minutes after active questioning and providing of correct answers on the occasion R_1. Of most interest is the effect of sleep loss on delayed recall.

Delayed recall during sleep loss (score R_3 minus score R_2) is shown in Fig. 13. Recall after 24 hr. (R_3) is as accurate as recall after a few minutes (R_2) for the first day of sleep loss. After 51 hr. of sleep loss have gone by, however, delayed recall is markedly poorer and remains so until the recovery period. Thus, this loss does not increase steadily as sleep loss increases (K = 0), but scores are, in general, lower during sleep loss than during baseline and recovery periods (T = .70*). This change in sleep loss cannot be accounted for by a lapse hypothesis. If the results could be substantiated, an additional concept would be necessary. Unfortunately, we have only the one task in which a change in delayed recall could appear.

[13] In all of the statistical analyses we used the difference between the scores of the experimental *S* and his matched control. This procedure compensates for practice effects and differential difficulty of lists.

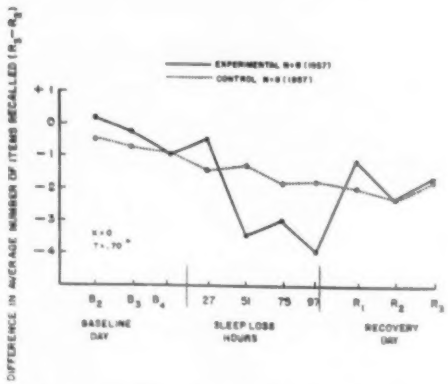

FIG. 13. The effect of sleep loss on delayed recall. [United States Army photograph]

Summary and Comment on E-*Paced Tasks*

It may seem strange that we put together a communication task, a learning task, and a number of vigilance tasks, without regard for the mental processes assumed to be involved. In all of these tasks, however, the effective response time is brief and controlled by *E*, and in all of them the results substantiate the expectation that in *E*-paced tasks impairment with sleep loss will take the form of an increase in errors. For *S*-paced tasks, it will be recalled, changes in speed rather than in accuracy were expected.

The incidence of errors was found to be mostly affected by the length of time during which a task continued without interruption and without a change in the stimulus response conditions. Task duration in this sense interacts with sleep loss, so that with increasing hours of sleep loss, errors start to occur earlier and earlier in the task. The other condition considered—knowledge of results—showed only a slight and somewhat inconsistent effect on performance. These results are the same as those found with *S*-paced tasks.

At this point, we feel that the usefulness of the lapse hypothesis has been amply demonstrated. It offers a very fruitful pointer to the aspects of performance in which impairment will appear, and is a guide to the scores or measures that will be most sensitive to the effects of sleep loss. In addition, the occurrence of lapses in three sensory modalities and a large variety of tasks suggests that lapses are a central rather than a peripheral phenomenon.

With this last point in mind, we turn to material on changes in the EEG alpha

rhythm with sleep loss and to relationships between these changes and the occurrence of lapses. This material is particularly relevant to the identification of lapses as transient states of sleep or a condition close to sleep.

An Electro-Physiological Correlate of Impairment[14]

In the history of sleep loss, there have been a number of attempts to find impairment in physiological functions. Of these, the most consistent positive result is a decline in the EEG alpha rhythm with eyes closed (Bjerner, 1949; Kleitman, 1939; Malmo, 1958; Tyler et al., 1947). Changes in alpha rhythm also have been shown to be related to brief changes in performance.

In Bjerner's 1949 study of sleep loss, for example, depression of alpha activity was significantly associated with "delayed actions" on a reaction-time task. In a later study, Simon and Emmons (1956) found a supporting result with sleepy or sleeping *S*s: as the quality and quantity of alpha increased, so did the probability that a stimulus would be reported heard and correctly recalled later. The stimuli were answers to information questions.

These findings strongly suggest that changes in alpha rhythm are a likely place to find a physiological correlate of impairment and lapses. In this section we present part of the material gathered by Armington and Mitnick (1959) substantiating a general decline in alpha with sleepiness, together with the results of an auditory vigilance task with simultaneous EEG recording. On this task, errors of omission were significantly associated with lower alpha activity. It will be recalled that, in an *E*-paced task like vigilance, errors of omission reflect the occurrence of lapses. We are strongly inclined to follow Bjerner's argument that both the general decline in alpha with sleep loss and the more transient declines accompanying brief impairments of performance signal a state that is similar to sleep itself.

Sleep loss and a decline in alpha amplitude. As a start, let us first establish that in the present study, alpha with eyes closed declines with sleep loss. The decline is clearly shown in Fig. 14, taken from work by Armington and Mitnick (1959).[15]

In Fig. 14, alpha during sleep loss is considerably less when the *S* is instructed to "keep his mind blank" or to count from one to ten than when he is instructed to add 17's. This is a reversal of the relationship found in the normal awake state (control period). A similar reversal has been noted before with fatigue (see Bjerner, 1949, p. 74) and with narcolepsy (Daly & Yoss, 1957). The result gives strong support to the frequently advanced hypothesis that a decline in alpha occurs with a shift, in either direction, away from an intermediate level of alertness, tonus, excitation, or mental activity (cf. Bjerner, 1949; Lindsley, 1951; Stennett, 1957). In a normal awake state, a shift upwards and a depression of alpha can be induced by mental work or by a startle stimulus. During sleep loss, the depression of alpha is induced by a shift down from the intermediate level; in this state the introduction of mental work or a startle stimulus shifts the *S* away from a drowsy state, back toward the intermediate level of activity and the reappearance of alpha.

Alpha and errors of omission. We wish to compare the presence of alpha during errors of omission with the amount present when other responses are made. With this aim in mind, the response to each stimulus on the auditory vigilance task was classified as follows: (*a*) correct response—the critical signal (the letter X) occurred, and *S*

[14] For the material in this section, the guidance and help of John Armington was invaluable.

[15] For Fig. 14 and all subsequent results in this section, scores were based on an automatic scoring device developed by Armington, Biersdorf, and Mailloux (1958). This device marks the presence of all 8–12 CPS activity with amplitude above a fixed voltage. For eight *S*s who were selected for high-amplitude alpha activity, the scoring level was set at 25 microvolts. For the other two *S*s, selected for low alpha, the level was set at 15 microvolts. For all *S*s, bipolar recording was used, with one electrode on the vertex and the other on the occipital prominence.

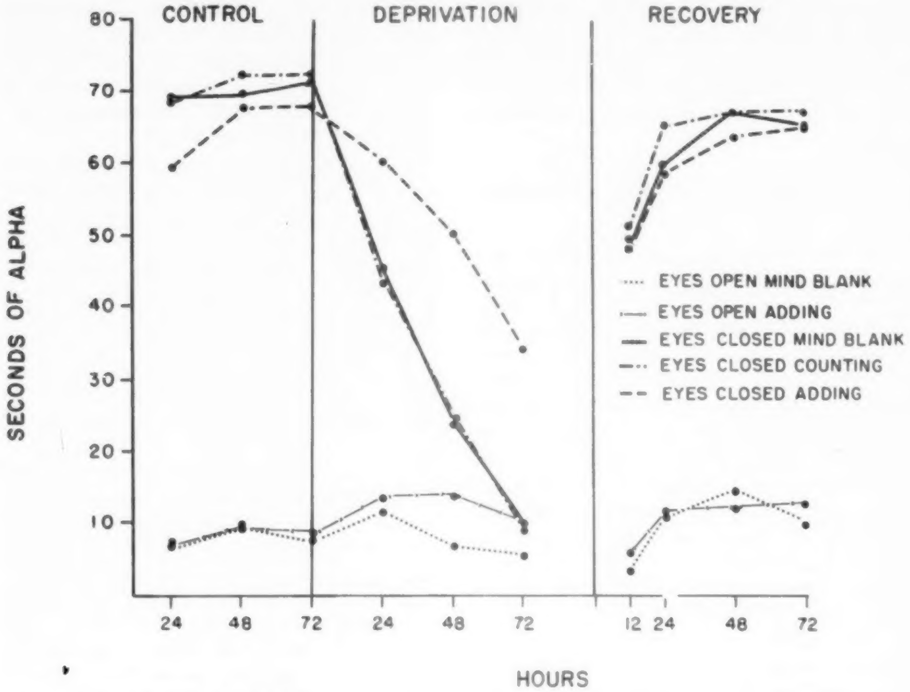

FIG. 14. Changes in alpha under several task conditions. [United States Army photograph]

responded with an appropriate release of the lever; (*b*) correct withholding of response—a letter other than X occurred, and *S* did not release the lever; (*c*) error of commission—the *S* released the lever for a letter other than X; (*d*) error of omission —the letter X occurred, and *S* failed to release the lever.

In determining the presence of alpha for each of these four types of response, the immediate problem is one of determining the time period around each response during which alpha could be counted. Simon and Emmons (1956) used the 2–7 sec. interval (average was 4 sec.) during which the answer to the information question was given to *S*. Bjerner (1949) used the time during which the delayed action occurred (most of these were 2–3 sec.), together with the 3 sec. before and after each delayed action. We have followed Bjerner in using a before-during-and-after interval, but have narrowed the interval down to 1 sec. before and 1 sec. after each stimulus letter. In our vigilance task, stimuli occur at the rate of one letter per second, and a longer interval gives rise to the problem that a single interval sometimes contained two errors of omission.[16]

For each of these brief intervals around each stimulus and response, alpha was scored as present or absent. Alpha was scored as present if there was any deflection of the automatic scoring pen (to avoid response to noise, the pen's action was so de-

[16] For one *S*, as a trial procedure, we used three time intervals to compare differences in alpha activity for the several responses. These intervals were: (*a*) from 1 sec. before to 1 sec. after the stimulus (results significant at .001 level), (*b*) over the 4 sec. before a stimulus (significant at .05 level), and (*c*) over the 2 sec. before a stimulus (significant at .01 level).

layed that a short burst of alpha was required to deflect it).

For 10 experimental *S*s in the 1957 group, Fig. 15 shows the average incidence of alpha for each of the four kinds of response that can accompany a stimulus. Errors of omission consistently exhibit the lowest amount of alpha activity, and the difference between alpha present for errors of omission, and the amount present with the other responses appears to increase with increasing sleep loss. The alpha scores return to baseline levels after 24 hr. of recovery.

Interpreting a decline in alpha. How are we to interpret both the over-all decline in alpha and the coincidence of lower alpha activity with errors of omission? If we follow Kleitman (1939) or Bjerner (1949), the over-all decline signals a state that is close to sleep. For Bjerner, a lapse is explained in the same way. It "is a transient condition of the same nature as sleep" (p. 73).

Bjerner bases his arguments on three points: (*a*) alpha activity is known to decline and disappear with the onset of sleep. (*b*) In long delayed actions (over 5 sec.), "all of the electroencephalographic changes considered to be characteristic of sleep occurred to a greater or less degree" (p. 68), e.g., slow potentials and K-complexes in response to sound stimuli. Furthermore, on a number of long delayed actions, the *S*s were observed to in fact sleep: they showed "immobility, sometimes open mouth and snoring and no reaction to weak sounds but awakening on loud sounds or shaking" (p. 68). (*c*) In a study by Davis, Davis, Harvey, and Hobart (1938), sleepy *S*s were in-

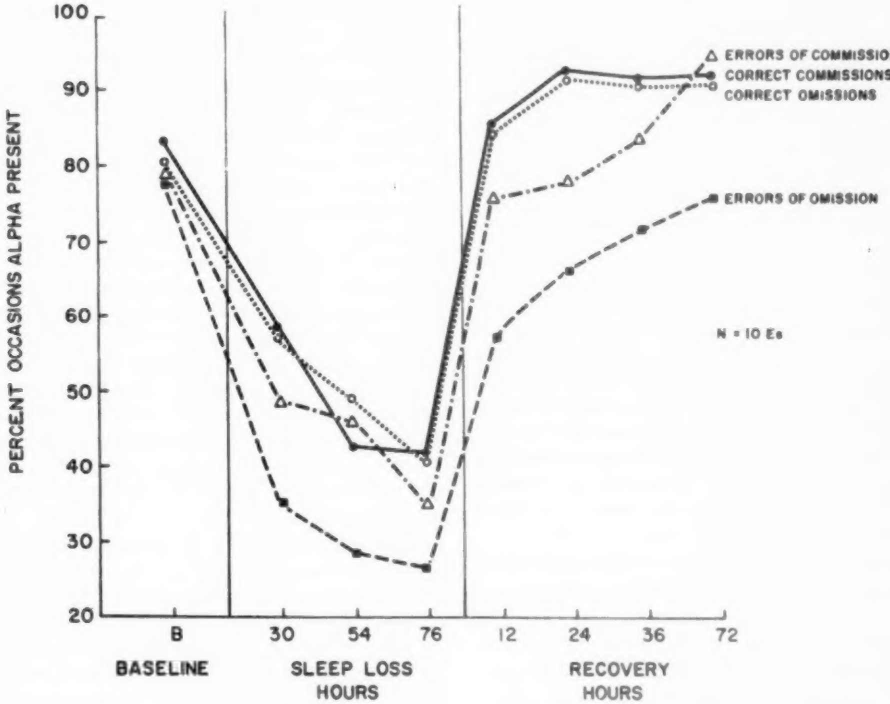

FIG. 15. The incidence of alpha with errors and correct responses on a vigilance task. [United States Army photograph]

structed to give a signal whenever they felt a spell of drowsiness. Most of these signals were preceded by depressions of alpha.

Additional support for regarding the decline in alpha amplitude as a sign of a condition like sleep was found in the present study by Armington and Mitnick (1959). The wave-form of the EEG was observed to be similar to that described by Simon and Emmons (1956) and by Jung (1954) as typical of drowsiness or of the transition state between wakefulness and sleep. On this basis, the decline would be expected to be greatest on a task which facilitates drowsiness ("mind blank" or counting from one to ten), and less on a task which arrests drowsiness (for example, adding 17's). These are, in fact the results presented in Fig. 14.

Although these various results present a consistent picture of changes in alpha occurring with a condition like sleep, it must be noted that an alternative hypothesis has been proposed. This is the hypothesis that the decline of alpha reflects the *S*'s increasing efforts to stay awake and to maintain performance (Malmo, 1958; Tyler et al., 1947). The hypothesis is limited to the over-all decline in alpha; i.e., it does not consider the transient declines occurring with lapses.

It is difficult to see how a hypothesis of increased effort could account for the reversal, during sleep loss, of the usual relationship between alpha depression and mental activity (Fig. 14). During sleep loss there is more alpha present when the *S* is adding 17's than when he is simply counting from one to ten. It appears more plausible to us to interpret the difference between tasks as indicative of the extent to which a task facilitates or arrests a state of drowsiness or near-sleep.

Summary. A decline in EEG alpha amplitude was observed with increasing sleep loss, a result which substantiates previous work. In addition, errors of omission on an auditory vigilance task were consistently associated with less alpha than were other responses. Both of these findings are interpreted as indicating that in sleep loss generally, and during lapses particularly, the *S* is in a state similar to or close to sleep itself.

Discussion

With the results and the ideas presented, we had hoped: (*a*) to account for the apparently inconsistent effects of sleep loss, (*b*) to provide some guide toward effective testing for impairment, and (*c*) to make a start on describing the general nature of impairment, and relating psychological to physiological changes.

As a way of accounting for negative and inconsistent results, the relationships between lapses and the properties of a task are decidedly useful. They enable us to draw together within the present study the results of a large variety of tests. Moreover, when we look back at previous studies, negative results can often be explained in terms of failure to measure the sensitive aspect of performance. The apparent lack of effect upon difficult problem-solving tests, for example, is largely attributable to measuring accuracy but not speed on *S*-paced tasks (Edwards' [1941] untimed intelligence test, for example). Where, in fact, both aspects have been measured, there are a number of results which point toward the importance of time. Weiskotten and Ferguson (1930) found changes in speed but not in accuracy on a task of mental arithmetic. Carmichael, Kennedy, and Mead (1949) observed a decline in reading speed but not in reading comprehension. And Laslett, using an intelligence test, commented that "Since many of them [the *S*s] had a sufficient amount of time to check their work on the test once or even twice, . . . the losses shown by these subjects . . . are less than the true losses" (Laslett, 1928, p. 391).

In a similar vein, the inconsistencies surrounding attempts to vary the *S*'s interest or knowledge of results can be at least partially resolved by the general condition that Bartlett (1942) has stressed: namely, the extent to which the *S* relaxes his standards of performance. With this general condition, attention focuses on the way in which any particular introduction of information, re-

minders, warnings, or exhortations counteracts the relaxation of standards, and shifts away from a provocative but rather unwieldy condition like motivation or arousal (Ax et al., 1957; Wilkinson, 1957).

To take the second aim, how effective a guide in testing for impairment are the ideas and results presented? In the first place, they run counter to the arguments that, since test performance shows so little change, we should concentrate on changes in behavior and mood (Kleitman, 1939) or on the measurement of effort (Tufts College, 1949). In fact, it now seems doubtful to us that the much-discussed changes in mood and disposition (Kleitman, 1939; Tyler et al., 1947) are directly attributable to sleep loss, either by way of a loss of judgment or a lowered threshold of irritability. In the present study and in the study by Eagles et al. (1953), incidents of anger and irritability were rare. In both studies, when anger appeared it was in resentment of a particular method of being kept awake or a face-saving denial of having fallen asleep. It seems quite likely that disposition changes are primarily a function of the *S*'s attitudes and the methods of control rather than any direct effect of sleep loss.

More specifically, the relationships between lapses and the properties of a task provide some very definite guides to the aspects and measures of performance which are most likely to be sensitive. In this respect, there is some guarantee against the conclusion of "no impairment" which often follows a failure to measure the critical aspect. Furthermore, the conditions found to promote impairment—task monotony and the opportunity to relax standards—suggest ways of making the appearance of impairment more likely.

There is, however, a major gap in the concepts presented. This lies in the relationship between lapses and performance on tasks which are more automatic and less dependent on stimulus discrimination than the ones we have used. Singleton (1953), for example, using a task where perceptual and motor speed could be measured separately, found that only the perceptual part of performance was affected. And Wilkinson (1958), using a vigilance task where the lapse hypothesis would lead one to concentrate on errors of omission or nonresponse to a stimulus, found that errors of commission become more important if a rhythmic sequence of responses was set up.

The difficulty in handling these tasks seems to lie, as Broadbent (1953b) has suggested, in the definition of a lapse as a period of no response. It may be more appropriate to regard lapses as periods in which the *S* is less responsive to a stimulus, and in which the gap may be appropriately filled by a response which is highly practiced and less dependent on stimulus discrimination, or inappropriately filled by a competing response. This type of definition would fit with the analyses of motor tasks that have stressed the disorganization or dislocation of units in performance (Bartley & Chute, 1947; Welford, 1953), and would also fit with Bills' (1935, 1937) finding that blocks are promoted by the presence of competing responses and lessened by practice. Some such additional definition would clearly be needed before the lapse hypothesis could be used with profit for tasks where the response hinges less upon the discrimination of a stimulus. For the tasks used in the present study, however, the definition of a lapse as a period of no response is adequate.

The final question, and the most difficult, is the general nature of impairment. Up to a point, it is extremely helpful to view impairment during sleep loss in terms of lapses which reflect brief periods of sleep or a condition like sleep. This approach, for example, makes it possible to obtain very clear relationships between performance on psychological tests and changes in alpha rhythm.

The need to extend this view of impairment comes from two points: (*a*) there is as yet no clear relationship between the nature of impairment and the conditions promoting impairment, and (*b*) lapses have been found to characterize impairment under conditions which do not appear to be the same as sleep loss, namely hypoxia (Maag, 1957) and psychosis (Eysenck,

Granger, & Brengelmann, 1957, pp. 100–101).

On the first of these two points, we can do little more than point to the problem and some suggested solutions. At an empirical level, for example, there is no difficulty in stating that impairment is promoted by an uninterrupted and unchanging set of stimuli and responses, both in studies of sleep loss and in studies of fatigue. The problem lies in describing this effect in less empirical terms.

It is this problem which is the core of Solomon's (1948) emphasis on stimulus satiation in prolonged work, and of Broadbent's (1957) model for perception and memory in terms of stimulus satiation and stimulus sensitivity. It is also the heart of Kleitman's (1939) view of wakefulness as maintained by incoming stimuli and accordingly decreased either by monotonous stimuli or by the lesser stimulation that goes with physical immobility. None of these approaches offers a very specific guide as to the optimal arrangement of stimuli and responses, but the proposals of Solomon (1948) and Broadbent (1957) do suggest some possible and needed lines of research.

What are we to make of the finding that lapses characterize impairment under conditions like hypoxia (Maag, 1957) and psychosis (Eysenck et al., 1957)? These results make it impossible to regard lapses as always reflecting brief periods of a condition like sleep. What appears far more likely, and indeed more encouraging, is that impairment under any condition takes the form of an increasing irregularity or unevenness of performance (cf. Maag, 1957). For the specific condition of sleep loss, the major contributor to the unevenness is the occurrence of brief periods of sleep. From this point of view, the contributing factors may vary from one impairing condition to another, but performance itself always contains the possibility of unevenness. This feature of performance has been well described by Broadbent: "Crudely speaking, a man is not like a child's mechanical toy which goes slower as it runs down. Nor is he like a car engine which continues normally until its fuel is exhausted and then stops dead. He is like a motor which after much use misfires, runs normally for awhile, then falters again, and so on" (Broadbent, 1955, p. 2). This description is as apt for performance under sleep loss as it is for performance with prolonged work.

Summary

With 49 *S*s, deprived of sleep for 72–98 hr., performance deteriorated on a variety of tasks, an unusual result in studies of sleep loss. Deficit took the form of lapses (brief periods of no response accompanied by extreme drowsiness and a decline in EEG alpha amplitude). Four features of lapses were noted. (*a*) They occur in other conditions such as fatigue and hypoxia and appear to characterize impairment in general. (*b*) They increase in both frequency and duration as sleep loss progresses. (*c*) They are strongly affected by stimulus monotony. (*d*) Their specific effect on performance varies with the properties of the task. In *S*-paced tasks, for example, speed is the critical measure; in *E*-paced tasks, errors are critical. To identify the sensitive aspect of performance becomes the crucial problem.

REFERENCES

Armington, J. C., Biersdorf, W. R., & Mailloux, L. D. An electroencephalogram recording apparatus. *Amer. J. Psychol.*, 1958, **71**, 594–599.

Armington, J. D., & Mitnick, L. L. Electroencephalogram and sleep deprivation. *J. appl. Physiol.* 1959, **14**, 247–250.

Ax, A. F., Fordyce, W., Loovas, I., Meredith, W., Pirojnikoff, L., Shmavonion, B., & Wendahl, R. Quantitative effects of sleep deprivation. *USA Qm. Res. Develpm. Res. Rep.*, 1957.

Bartlett, F. C. Fatigue following highly skilled work. *Proc. Roy. Soc.*, 1942, **131**, 247–257.

Bartley, S. H. Fatigue and inadequacy. *Physiol. Rev.*, 1957, **37**, 301–324.

Bartley, S. H. & Chute, E. *Fatigue and impairment in man.* New York: McGraw-Hill, 1947.

Bills, A. G. Blocking: A new principle in mental fatigue. *Amer. J. Psychol.*, 1931, **43**, 230–245.

Bills, A. G. Some causal factors in mental blocking. *J. exp. Psychol.*, 1935, **18**, 172–185.

Bills, A. G. Fatigue in mental work. *Physiol. Rev.*, 1937, **17**, 436–453.

Bjerner, B. Alpha depression and lowered pulse rate during delayed actions in a serial reaction test: A study in sleep deprivation. *Acta Physiol. Scandinavica*, 1949, **19**, Suppl. 65. Pp. 1–93.

Broadbent, D. E. Neglect of the surroundings in relation to fatigue decrements in output. In W. F. Floyd & A. T. Welford (Eds.), *Fatigue.* London: Lewis, 1953. Pp. 173–178. (a)

Broadbent, D. E. Noise, paced performance and vigilance tasks. *Brit. J. Psychol.*, 1953, **44**, 295–303. (b)

Broadbent, D. E. Variations in performance arising from continuous work. In *Conference on individual efficiency in industry.* Cambridge, England: Med. Res. Council, 1955. Pp. 1–5.

Broadbent, D. E. A mechanical model for human attention and immediate memory. *Psychol. Rev.*, 1957, **64**, 205–215.

Bruner, J. S., Goodnow, Jacqueline J., & Austin, G. A. *A study of thinking.* New York: Wiley, 1956.

Carmichael, L., Kennedy, J. L., & Mead, L. C. Some recent approaches to the experimental study of fatigue. *Proc. Nat. Acad. Sci.*, 1949, **35**, 691–696.

Daly, D. D., & Yoss, R. E. Electroencephalogram in narcolepsy. *EEG clin. Neurophysiol.*, 1957, **9**, 109–120.

Davis, H., Davis, P. A., Harvey, E. N., & Hobart, G. Human brain potentials during the onset of sleep. *J. Neurophysiol.*, 1938, **1**, 24–38.

Eagles, J. B., Halliday, A. M., & Redfearn, J. W. T. The effect of fatigue on tremor. In W. F. Floyd & A. T. Welford (Eds.), *Fatigue.* London: Lewis, 1953. Pp. 41–58.

Edwards, A. S. Effects of the loss of 100 hours of sleep. *Amer. J. Psychol.*, 1941, **54**, 80–91.

Eysenck, H. J., Granger, G. W., & Brengelmann, J. C. *Perceptual processes and mental illness.* New York: Basic Books, 1957.

Goodnow, Jacqueline J. Determinants of choice-distribution in two-choice situations. *Amer. J. Psychol.*, 1955, **68**, 106–116.

Goodnow, Jacqueline J., Rubenstein, I., & Shanks, Betty L. The role of past events in problem solving. *J. exp. Psychol.*, in press.

Hyman, R., & Jenkin, N. S. Involvement and set as determinants of behavioral stereotypy. *Psychol. Rep.*, 1956, **2**, 131–146. Monogr. Suppl. 3.

Jonckheere, A. R. A distribution-free K-sample test against ordered alternatives. *Biometrika*, 1954, **41**, 133–145.

Jung, R. Correlation of bioelectrical and autonomic phenomena with alterations of consciousness and arousal in man. In J. F. Delafresnaye (Ed.), *Brain mechanisms and consciousness.* Oxford: Blackwell, 1954. Pp. 310–339.

Kleitman, N. *Sleep and wakefulness.* Chicago: Univer. Chicago Press, 1939.

Laslett, H. R. An experiment on the effects of loss of sleep. *J. exp. Psychol.*, 1924, **7**, 45–58.

Laslett, H. R. Experiments on the effects of the loss of sleep. *J. exp. Psychol.*, 1928, **11**, 370–396.

Lee, M. A., & Kleitman, N. Studies on the physiology of sleep: II. Attempts to demonstrate functional changes in the nervous system during experimental insomnia. *Amer. J. Physiol.*, 1923, **67**, 141–151.

Lindsley, D. B. *Emotion.* In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 473–516.

Lyerly, S. B. The average Spearman rank correlation coefficient. *Psychometrika*, 1952, **17**, 421–428.

Maag, C. H. Characteristics of mental impairment in hypoxia. *Amer. J. Psychol.*, 1957, **70**, 243–247.

Mackworth, N. H. Researches on the measurement of human performance. *MRC Spec. Rep. No. 268.* London: Her Majesty's Stationery Office, 1950.

Malmo, R. B. Certain physiological correlates of psychomotor functioning. Office of the Surgeon General, *USA Res. Rep.*, Contract No. DA-49-007-MD-626, 1958.

Murray, E. J., Schein, E. H., Erikson, K. T., Hill, W. F., & Cohen, M. The effects of sleep deprivation on social behavior. *J. soc. Psychol.*, 1959, **49**, 229–236.

Murray, E. J., Williams, H. L., & Lubin, A. Body temperature and psychological ratings during sleep deprivation. *J. exp. Psychol.*, 1958, **56**, 271–273.

PATRICK, G. T. W., & GILBERT, J. A. On the effects of loss of sleep. *Psychol. Rev.*, 1896, **3**, 469–483.

RODNICK, E. H., & SHAKOW, D. Set in the schizophrenic as measured by a composite reaction time index. *Amer. J. Psychiat.*, 1940, **97**, 214–225.

ROSVOLD, H. E., MIRSKY, A. F., SARASON, I., BRANSOME, E. D., JR., & BECK, L. H. A continuous performance test of brain damage. *J. consult. Psychol.*, 1956, **20**, 343–350.

SCHEIN, E. H. The effects of sleep deprivation on performance in a simulated communication task. *J. appl. Psychol.*, 1957, **41**, 247–252.

SIMON, C. W., & EMMONS, W. H. EEG, consciousness, and sleep. *Science*, 1956, **124**, 1066–1069.

SINGLETON, W. T. Deterioration of performance on a short-term perceptual-motor task. In W. F. Floyd & A. T. Welford (Eds.), *Fatigue*. London: Lewis, 1953. Pp. 163–172.

SOLOMON, R. L. The influence of work on behavior. *Psychol. Bull.*, 1948, **45**, 1–40.

STENNETT, R. G. The relationship of alpha amplitude to the level of palmar conductance. *EEG clin. Neurophysiol.*, 1957, **9**, 131–138.

TUFTS COLLEGE, INSTITUTE FOR APPLIED EXPERIMENTAL PSYCHOLOGY. *Handbook of human engineering data.* Pt. 7. Medford, Mass.: Tufts Coll., 1949.

TYLER, D. B. Fatigue of prolonged wakefulness. *Nat. Res. Council Res. Rep.*, 1946, OE Memr-282, Part 2.

TYLER, D. B., GOODMAN, J., & ROTHMAN, T. The effect of experimental insomnia on the rate of potential changes in the brain. *Amer. J. Physiol.*, 1947, **149**, 185–193.

WARREN, N., & CLARK, B. Blocking in mental and motor tasks during a 65-hour vigil. *J. exp. Psychol.*, 1937, **21**, 97–105.

WEISKOTTEN, T. F., & FERGUSON, J. E. A further study of the effects of loss of sleep. *J. exp. Psychol.*, 1930, **13**, 247–266.

WELFORD, A. T. The psychologist's problem in measuring fatigue. In W. F. Floyd & A. T. Welford (Eds.), *Fatigue*. London: Lewis, 1953. Pp. 183–192.

WILKINSON, R. T. Effects of lack of sleep. Cambridge, England: Med. Res. Council, *FPRC Rep. No. 961.3*, 1957.

WILKINSON, R. T. The effects of sleep loss on performance. Cambridge, England: Med. Res. Council, *APRU Res. Rep. No. 323*, 1958.

(Received October 11, 1958)